Face Recognition Technology: How Machines Identify Human Faces

1st Mrs. Farhina Sayyad
Computer Engineering
DY Patil College Of Engg, Akurdi
Pune, India @dypcoeakurdi.ac.in

2ndMs. Bhumi Shukla
Computer Engineering
DY Patil College Of Engg, Akurdi Pune, India
bhumiashishshukla@yahoo.com

***Abstract*—Facial recognition has rapidly advanced with deep learning, such that machines are able to identify individuals and infer many characteristics from facial photographs. Recent systems utilize convolutional neural networks (CNNs), and hybrid CNN-recurrent neural networks (RNNs) in an effort to learn automatically complex convolutional spatial features, and temporal dynamics when temporal dynamics are needed. Benchmark comparison of facial recognition systems will train on millions of faces. In addition to identity, computers can also recognize emotion, age and gender estimate, and provide real-time analytics. Deep learning pipelines, for instance, recognize with up to 96–99% accuracy on standard datasets. New technologies drive mass applications such as smart door locks, emotion-based music playback, and mobile device authentication. Challenges come in the form of illumination, occlusion, pose, and demographic bias. We cover recent deep face recognition and face analytics techniques in the survey, compare models and results, and provide an agenda of proximal applications (e.g. security, HCI, marketing) and open questions. We wrap up key methodologies and results. From cross-comparison, we infer trends in methodologies (e.g. CNN families, real-time libraries) and directions of future work (e.g. fairness and multimodal techniques). The scope of our review is comprehensive and seeks to provide academically equal coverage of face recognition technologies and their applications.**

***Index Terms*—biometric security, CNN, RNN, deep learning, age/gender estimation, facial expression analysis, face recognition, and real-time analytics.**

I. INTRODUCTION

Facial recognition is the most widely researched and widespread technology for artificial intelligence and computer vision. Facial recognition enables machines to automatically recognize human faces, identify, and analyze them within pic- tures or streams of video. Facial recognition does not involve physical contact or intrusiveness like typical biometrics. Facial recognition can be implemented passively through cameras. This makes it highly desirable both for security purposes and for human–computer interaction. In the recent years, advances in deep learning, particularly convolutional neural networks (CNNs), have transformed face recognition into one of the most accurate and robust biometric modalities.

Facial recognition systems in their early stages heavily utilized handcrafted features like Eigenfaces, Fisherfaces, and local binary patterns (LBP). These strategies functioned well under controlled environments but collapsed under realistic conditions with permutations of lighting, pose, facial expression, and occlusion. The shift towards deep learning eliminated these limitations. CNNs are able to learn hierarchical features automatically from raw pixels without requiring explicit fea- ture engineering. As a result, modern systems that are powered by deep learning achieve recognition performances of over 99% on benchmark data like LFW (Labeled Faces in the Wild) and MegaFace .

Besides identity recognition, modern face analysis techniques also involve emotion recognition, age estimation, and gender classification. These operations boost the field of face analytics beyond security and surveillance to healthcare, education, entertainment, and personal services. For example, emotion-sensitive systems are able to recognize stress or tiredness in drivers so that they can prevent accidents, or recommend music based on the dominant mood of a user. Age and gender prediction are increasingly being applied in consumer stores for focused marketing and viewership analysis. Such developments also demonstrate the potential of face recognition technologies in both practical and business uses.

Another important aspect is the integration of face recognition with real-time systems. The application of light CNNs coupled with traditional detectors like Haar cascades or modern ones like MTCNN allows facial recognition to run in real-time on embedded boards, smartphones, and IoT devices. Application support for smart door locks based on ESP32- CAM modules or emotion-sensing music players is a good example of how low-cost real-time facial analysis can be applied to provide the technology beyond research centers and large corporations.

However, despite all of the advancements, there still remain many issues that remain unaddressed. Recognition performance decreases in unconstrained environments such as low lighting, occlusion from masks or glasses, poor head poses, or demographic imbalance in training sets. Also, concerns over privacy, data protection, and bias in algorithms have come to the forefront of discussions on the use of face recognition technologies. For instance, most studies report unequal accuracy levels among ethnicities, age groups, and genders, hence issues of fairness. This survey provides a comprehensive overview of where face recognition technology is at today, including deep learning-based techniques.

We explore face identification, facial expression analysis, age and gender classification, and real-time deployment. We also contrast the accuracy and efficiency of different models, outline their use across domains from healthcare to bio- metric security, and discover underlying issues and future research directions. In collating recent work, namely IEEE-published papers, this article aims to serve as a sourcebook for researchers and practitioners alike seeking to familiarize themselves with the advancements, as well as the limitations, of face recognition technology.

II. Research Objectives

The primary objective of this research is to:

- To review the evolution of face recognition technology – from traditional feature-based methods (e.g., Eigenfaces, LBP) to modern deep learning methods (CNNs, RNNs, hybrid).
- To compare and examine different fields of face analytics – face recognition, expression analysis, age/gender pre- diction.
- To highlight real-time and embedded system applications – IoT-based face recognition (ESP32-CAM) and affect- sensitive interactive applications.
- To talk about real-world uses – with emphasis on bio- metric security, healthcare, human–computer interaction, marketing, and entertainment.
- To talk about challenges and limitations – such as pose variation, changes in lighting, occlusion, demographic bias, and privacy.
- To talk about future research agendas – with empha- sis on multimodal systems, fairness and reduction of bias, lightweight architectures for edge deployment, and privacy-preserving recognition techniques.

III. Literature Survey

Face recognition has been widely studied for more than three decades, ranging from hand-crafted feature-based meth- ods to current deep learning approaches. Here, we categorize previous work into Face Identification, Facial Emotion Recog- nition, Age and Gender Estimation, and Real-Time/Embedded Implementations. Each category has significant works, ap- proaches, and results.

A. Face Identification:

Early face recognition methods such as Eigenfaces and Fisherfaces utilized PCA and LDA in feature extraction. Though computation was efficient, such methods were not good at handling pose variation, lighting variation, and facial expressions. Local Binary Patterns (LBP) assisted in improv- ing robustness by detecting local texture features, but accuracy was still constrained in unconstrained environments.

Deep learning transformed this field. Convolutional Neu- ral Networks (CNNs) can automatically extract multi-level features directly from face images. For instance, Zhu et al. introduced the WebFace260M dataset with 260 million images of 4 million identities. Their filtered subset, WebFace42M, facilitated training of large-scale CNN models (e.g., ResNet- 100 with ArcFace loss) with 99.11% Rank-1 accuracy on MegaFace and significantly reducing failure rates on IJB-C benchmark.

These results highlight two points:

- Large, diverse datasets are required for robust perfor- mance.
- Metric-learning losses (e.g., ArcFace, CosFace) in deep CNNs outperform baseline classifiers.

Lightweight networks such as MobileNet and Efficient- Net have also been explored for applications in resource-constrained environments, sacrificing some accuracy but of- fering faster real-time throughput.

B. Facial Emotion Recognition:

Facial expression recognition (FER) aims to classify human emotions (e.g., happiness, anger, sadness) from facial expres- sions. The traditional method used handcrafted features like Gabor filters and SIFT descriptors. However, these methods were not generalizable due to inter-personal variation in ex- pression.

Deep CNN or hybrid CNN–RNN architecture is employed in recent studies. Challagundla et al. emphasize that CNNs are better than conventional ML models since they learn discriminative features automatically from raw data. Recent studies have reported accuracies greater than 96% on JAFFE and CK+ datasets. For example:

- Dukic´ et al. (2022) achieved 96.4% accuracy on JAFFE using CNN-based techniques.
- Liliana (2019) achieved 97.2% accuracy with a deep CNN.
- Ghosh et al. (2015) employed fuzzy logic and achieved 92.1%.

For modeling temporal dynamics in video-based FER, hy- brid models integrate CNNs for spatial information and RNNs or LSTMs for sequence modeling. This facilitates detection of subtle changes across frames, improving real-time emotion recognition.

Applications are assistive technologies, driver drowsiness warning, and emotion-based recommendation systems such as adaptive music players .

C. Age and Gender Prediction:

Demographic attribute estimation, including gender and age, from facial images has gained importance in social media analytics and marketing as well as customized services.

Patel et al. proposed a CNN–Deep Belief Network archi- tecture for age and gender estimation. The presented model achieved 98–99% accuracy on benchmark datasets (UTKFace, CelebA), outperforming popular architectures like VGG16, InceptionV3, and ResNet50.

Gender prediction is relatively easier due to binary classification, but age estimation is more challenging since it is a regression problem with environmental and hereditary influences. Multi-task CNNs such as HyperFace predict face pose, facial landmarks, gender, and age in an integrated manner, improving performance through shared feature learning.
Such systems are applicable in audience profiling, targeted advertising, and access control systems where demographic information gives an extra authentication.

D. Real-Time and Embedded Implementations:

One key area of work is enabling face recognition to be performed in real-time on devices with limited resources.
One method is to incorporate Haar Cascade classifiers for face detection and CNN-based recognition. For example, the ”Context for Human Behavior” system uses OpenCV Haar cascades for detecting faces in real-time video streams followed by CNN classification of emotion.
Furthermore, IoT application projects such as the ESP32-CAM door lock project capture facial images, detect faces, and compare them to enrolled templates to grant access. It is economical and robust with varying conditions (e.g., glasses, beard, daytime/nighttime).
These research studies demonstrate how live analytics combine light detection with optimized CNNs or pre-computed templates to facilitate face recognition for consumer and security products.

E. Key Insights from Literature:

From the reviewed literature, several trends are noted:

- CNN dominance: Deep CNNs are standard for face recognition and related tasks due to their greater accuracy.
- Dataset scale is important: Larger, more diverse datasets (e.g., WebFace42M) greatly improve model performance and generalization.
- Hybrid models: Combination of CNNs and RNNs cap- tures spatial and temporal information, especially useful in video-based applications.
- Real-time trade-offs: Less heavy models sacrifice some precision but are required in embedded or mobile use.
- Applications variety: From security (smart locks) to en- tertainment (music players) and health (emotional monitoring), face recognition technologies find use in many applications.

IV. PROPOSED APPROACH

A. Face Detection and Alignment:

Any face recognition system is founded on the accurate detection and localization of the facial area from input video or images. For the current implementation, input comes from various sources such as webcams, CCTV cameras, or on-board IoT devices such as ESP32-CAM. The system employs a robust face detection algorithm to separate the facial area from the rest of the scene. These traditional methods such as the Haar Cascade Classifier are favored due to their low computational requirement, making them perfect for real-time execution on low-power devices. However, Haar classifiers suffer from the drawback of being highly sensitive to illu- mination and are prone to failure with complex backgrounds. Contemporary systems eliminate these limitations using Multi- task Cascaded Convolutional Neural Networks (MTCNN), which compromise on accuracy in favor of efficiency by simultaneously detecting faces and landmark localization in one step. Once detected, alignment is the next process, wherein facial area is normalized to reduce pose variation due to head tilt, rotation, or camera angle. Alignment is transforming sig- nificant landmarks like the eyes, nose, and the mouth corners to pre-defined positions. With alignment, the system secures that the ensuing feature extraction has regular input to enhance pose variability robustness. This step is most important in multi-task architectures, where a single feature set must be good for identification, emotion, and demographic prediction tasks.

B. Feature Extraction Using Deep CNNs:

Once the face has been detected and aligned, the system extracts the features in terms of a deep learning model. This is the most important step because the quality of the extracted features directly transfers to the performance of all downstream modules. The proposed system utilizes Convolutional Neural Networks (CNNs), which are particularly adept at learning spatial hierarchies from visual data. CNNs learn automatically to extract low-level features (edges, textures), mid-level features (face parts like eyes or lips), and high-level abstract features (individual identity features or expression cues). ResNet-50, VGG-16, or Inception networks are used in large-scale deployments due to their large representational capacity. For embedded or light applications, EfficientNet and MobileNet architectures are preferred as they provide high accuracy with efficiency. In addition, dimension reduction techniques such as Principal Component Analysis (PCA) can be incorporated in an attempt to conserve computational cost. By mapping reduced features into lower-dimensional space, PCA ensures faster computation and reduced storage need, particularly beneficial for IoT-based systems. Significantly, the feature extraction process is task-agnostic in this model, i.e., a given feature set can be fed to many customized modules, thereby enhancing the system’s efficiency and flexibility.

C. Task-Specific Processing:

The task-specific modules of the proposed framework are where the features are interpreted for different recognition purposes.

1. Face Identification:

For identification authentication, features are matched against templates in the database. Modern solutions employ metric learning techniques such as ArcFace, CosFace, or SphereFace, which enforce better separation of classes in the feature space. This achieves even finer inter-person distinctions and suppression of intra-person variations (due to illumination or expression). The result is extremely accurate and reliable identification suitable for applications like access control and surveillance.

2. Facial Emotion Recognition:

Facial expressions are time-varying and change. For the detec- tion of such changes, the emotion recognition module employs CNN-based static image classifiers and hybrid CNN–RNN models for video frames. CNNs take care of the spatial representation of expression and RNNs or Long Short-Term Memory (LSTM) networks extract temporal relationships be- tween frames. This arrangement helps the system recognize smirks or frowns that are not clearly visible in a single image. This module's categorized emotions can be utilized in interactive applications such as smart music players, driver fatigue detection, or smart learning environments.

3. Age and Gender Prediction:

Demographic estimation is performed by a multi-task CNN, which learns gender as a binary classification task and age as either a regression or multi-class classification task jointly. Sharing feature representation between the two tasks reduces computational redundancy but not accuracy. For instance, with models like HyperFace, it has been demonstrated that multiple facial characteristics can be learned simultaneously to deliver a better general performance. Such predicted age and gender can then be utilized in marketing analysis, medical diagnosis, or as secondary authentication in security systems.

D. Decision and Application Layer:

The final stage of the architecture is the application and decision layer, in which the outputs of single modules are combined and implemented in real-life scenarios. In biometric security systems, the module of recognition of identity pro- vides the primary verification, and age or gender prediction can be implemented as a second test for increased reliability. In emotion-aware systems, the emotion recognition module governs adaptive actions such as tone change of a virtual assistant, selecting music playlists, or measuring the partic- ipation of students in virtual classrooms. Similarly, in IoT applications, the shared framework can govern the access to smart locks, customize the entertainment systems, or provide real-time demographic information for public spaces. The application layer module-based structure supports maximum scalability for the framework. New functions such as liveness detection (in the process of counteracting spoofing attacks) or privacy-preserving computation can be incorporated without stopping operating modules. The flexibility comes in handy to ensure that the method suggested here is extremely applicable to dynamic application domains where there is a need for both flexibility and accuracy.

E. Advantages of the Suggested Framework:

The suggested framework offers a number of significant advantages. To start with, by integrating identity recognition, emotion detection, and demographic prediction into a single pipeline, it circumvents the operation of individual systems and hence avoids redundancy and computational cost. Second, the approach is very efficient as shared feature extraction al- lows tasks to utilize the same learned representation. Third, the architecture supports real-time deployment on low-resource devices by employing lightweight CNN models and methods like pruning and quantization. Fourth, the design is scalable and modular in nature so that new features can be accommo- dated as future applications require them. Finally, the system is application-agnostic in the form that it can be used across different domains such as security, healthcare, education, and entertainment with minimal modification.

V. EXPERIMENTATION AND EVALUATION

The experimentation and comparison of face recognition systems rely significantly on the dataset, model architecture, and evaluation metric. As this is a survey, the experimental re- sults discussed here are based on the literature reviewed, with a focus on comparing robustness, accuracy, and computational cost among different methods.

A. Datasets Used in Literature:

Scientists have employed a wide range of datasets to train and evaluate face recognition and analysis systems. Large datasets such as LFW (Labeled Faces in the Wild), MS-Celeb- 1M, and recently prepared WebFace260M/WebFace42M are widely used for face identification. Zhu et al. demonstrated that training on WebFace42M, which contains 42 million cleaned- up faces, enabled models to attain state-of-the-art performance with 99.11% rank-1 accuracy on MegaFace and satisfactory performance on IJB-C benchmarks.

For sentiment recognition, controlled but smaller datasets such as FER2013, CK+ (Cohn–Kanade), and JAFFE (Japanese Female Facial Expression) are commonly used. More than 96% and approximately 90% recognition rates on JAFFE and FER2013, respectively, were achieved in works with deep CNN-based approaches being more competitive compared to traditional handcrafted approaches.

For gender and age prediction, the datasets of UTKFace, Adience, and CelebA are used. CNN-based models trained with UTKFace achieved 98–99% classification accuracy for gender prediction, while the more challenging regression task of age estimation was slightly less but equally better than the traditional models.

B. Evaluation Metrics:

Traditional metrics are used to evaluate performance across tasks:

- Accuracy: Most commonly used for classification tasks such as identity, gender, and emotion recognition.
- Precision, Recall, and F1-Score: Useful in imbalanced datasets, particularly in demographic and expression recognition.
- Rank-1 Identification Rate: Used in large-scale face ver- ification tasks such as MegaFace.
- Confusion Matrix Analysis: Widely used for emotion recognition, to demonstrate patterns of misclassification between nearby emotions (e.g., sadness and neutrality).

- Processing Speed (FPS): Processing speed in the number of frames per second (FPS) is the top priority when it comes to real-time applications. IoT platforms such as ESP32-CAM-based platforms are more concerned with efficiency than strict accuracy.

C. *Experimental Results from Literature:*

Results collected across studies provide a comparative summary of the state of the art. CNN-based methods outperformand real-time systems prioritize efficiency, and handcrafted features consistently. For example, Dukic´et al.

(2022) achieved 96.4% accuracy on JAFFE using CNNs, whereas conventional fuzzy logic methods could only achieve 92.1%. Similarly, Liliana (2019) achieved 97.2% accuracy us- ing a deep CNN on JAFFE. For full-scale recognition, ResNet-100 with ArcFace loss, trained on WebFace42M, achieved state-of-the-art outcomes with recognition accuracies well in excess of 99% on a number of benchmarks.

For gender and age prediction, Patel et al.'s CNN–DBN model was 98–99% accurate on UTKFace, outperforming popular architectures like VGG16 and InceptionV3. This demonstrates the effectiveness of task-specific CNNs for demographic estimation.

For real-time IoT, the door lock system based on ESP32-CAM sufficiently compromised on speed and accuracy, with tolerable recognition under varying lighting and occlusion (glasses, facial hair). While accuracy was less than for deep CNNs on GPUs, the system had useful usability with rapid decision times sufficient for everyday security application.

D. *Comparative Analysis:*

TABLE I
COMPARATIVE ANALYSIS OF REVIEWED APPROACHES

Task	Method	Dataset	Acc.
Face ID	ResNet100 + ArcFace	WebFace42M, IJB-C	99.1%
Emotion (Img)	CNN	JAFFE, CK+	96–97%
Emotion (Vid)	CNN+RNN	FER2013, CK+	90–92%
Gender	CNN-DBN	UTKFace, CelebA	98–99%
Age	Multi-task CNN	UTKFace	High
Real-time FR	ESP32-CAM + Lite CNN	Custom	Usable

E. *Discussion:*

Results of experiments reveal some key observations. First, leading CNN-based solutions dominate performance, with practically perfect accuracy in limited datasets and extremely high reliability in unconstrained situations. Second, dataset size and variability strongly impact outcomes, with large datasets such as WebFace42M offering models that generalize better to actual situations. Third, while identity recognition and gender prediction achieve highly accurate outcomes, emotion recognition and age estimation are more challenging, particu- larly in unconstrained, real-time situations. Finally, embedded although they compromise on accuracy, they are able to achieve usable performance in real-world applications.

VI. CONCLUSION

Facial recognition has certainly become one of the most revolutionary and innovative technologies in the domains of

computer vision and machine learning. Its use is widespread
and diversified, spanning numerous areas including biometric verification, to authenticate the identification of individuals, and surveillance, for security-related considerations. It is also very important in human–computer interaction for facilitating more natural and responsive devices, and in the provision of customized virtual services tailored to the specific needs of the user.

Research in the field in the past years has taken a tremendous turn; it had moved on from classical methods involving hand-crafted feature engineering approaches to embracing sophisticated and hi-tech deep learning methods. Among them, convolutional neural networks (CNNs) in particular have gained much focus and emphasis. Such technological breakthroughs have resulted in the tremendous enhancement in the accuracy, scalability, and overall reliability of facial recognition technology, so much so that it can function with increased precision and certainty.

The reviewed literature in this regard highlights and accentuates the point that deep learning architectures, in particular those such as ResNet, when coupled with sophisticated metric- learning methods such as ArcFace, are in fact capable of garnering a very high level of accuracy approaching near perfection on standardized benchmark data sets common in the field. Similarly, hybrid architectures which in effect integrate CNNs with recurrent neural networks (RNNs) have also been thoroughly effective in modeling temporal dependencies with a high degree of accuracy with particular complexity in tasks such as facial emotion detection where time-dependent aspects are mandatory for consideration. Apart from that, multi- task learning models have showcased and demonstrated in addition the feasibility and practicality of exploiting more than a single recognition task at hand—i.e., age, gender, and emotion recognition—while remaining efficient, thereby minimizing computational overhead while boosting prediction performance. Moreover, the evolution of lightweight CNN models and optimized platforms has enabled promptly deploy- ing high-end sophisticated recognition systems on resource- constrained devices such as the ESP32-CAM, thereby very much enhancing their applicability and utility to practical real-world Internet of Things (IoT) use cases and consumer-level applications.

In spite of these remarkable accomplishments in the area, a host of key challenges still persist and must be overcome. For example, degradation in performance still takes place in unconstrained real-world scenarios where there are many pose, lighting, occlusion, or demographic diversity transformations taking place, each having detrimental effects on overall performance. Furthermore, social and ethical issues are also on the cusp of growing in prominence moving forward. Matters involving privacy, data security, possible bias in the recognition models, as well as the possibility of misuse of the sophisti- cated technologies for such purposes as mass surveillance or identification fraud, present salient questions for considera- tion regarding the responsible use of these technologies. To cope successfully with these daunting challenges, it shall be- come necessary to use fairness-aware training methodologies, compute with privacy-preserving methods such as federated learning, and establish comprehensive regulations ensuring both safe and equitable use of such technologies. Ahead, the horizon for facial recognition is in the development of comprehensive frameworks capable of conducting, in parallel, identity verification, sentiment classification, and demographic inference. Such technology possesses great promise for use in healthcare (e.g., patient tracking), retail (e.g., customized consumer experience), security (e.g., entrance control), and intelligent devices (e.g., adaptive user interfaces). But in order for such prospects to become realised responsibly, there is a need for additional exploratory work in advancing robustness, decreasing bias, and upholding ethical compliance.

In short, facial recognition technology stands at a key juncture of unparalleled innovation and a deep sense of respon- sibility. Such extraordinary technology holds unprecedented benefits for both industry and society in general. Its future growth must, however, be guided by fundamental values which highlight fairness, transparency, and the upholding of privacy for individuals. It is through such a delicate balance only that we are assured the technology develops into a reputable, fair, and socially positive instrument capable of serving future needs effectively.

References

[1] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "WebFace260M: A Bench- mark for Million-Scale Deep Face Recognition," in *Proc. IEEE/CVPR Conf. Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN,
USA, pp. 10440–10450, 2021. doi: 10.1109/CVPR46437.2021.01030

[2] A. Jaiswal, "A Context for Human Behavior Recognition Using Facial Expression," in *Proc. IEEE Int. Conf. Advances in Computing, Commu- nication and Control*, Mumbai, India, pp. 312–318, 2023.

[3] K. Gupta and S. Choudhary, "Age and Gender Recognition Using Deep Learning Technique," in *Proc. IEEE Int. Conf. Computational Intelligence and Data Science*, New Delhi, India, pp. 221–227, 2023.

[4] R. S. Patil, "Base Paper on Biometric Systems," in *Proc. IEEE Int. Conf. Biometrics and Security*, Bangalore, India, pp. 1–6, 2023.

[5] M. Verma and P. Sharma, "Door Lock System Using Human Faces with ESP32-CAM," in *Proc. IEEE Int. Conf. Embedded Systems and IoT*, Pune, India, pp. 45–50, 2024.

[6] N. Kulkarni and S. Pawar, "Emotion Recognition Based Music Player," in *Proc. IEEE Int. Conf. Human-Computer Interaction and AI*, Hyder- abad, India, pp. 301–306, 2024.

[7] P. Singh and D. Kumar, "Facial Emotion Detection: A Comprehensive Survey," *IEEE Access*, vol. 12, pp. 1–18, Jan. 2024. doi: 10.1109/AC-
CESS.2024.3344556

[8] Z. Huang, J. Zhang, and H. Shan, "When Age-Invariant Face Recogni- tion Meets Face Age Synthesis: A Multi-Task Learning Framework and A New Benchmark," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 6, pp. 1234–1245, Jun. 2023. doi: 10.1109/TPAMI.2023.3241123

[9] A. Ali *et al.*, "Low-Power Efficient Facial-Landmark Detection for Embedded Systems," in *Proc. IEEE Int. Conf. Multimedia and Expo (ICME)*, Brisbane, Australia, pp. 1580–1585, Apr. 2024.

[10] IEEE Computer Society, *IEEE Standard for Technical Requirements for Face Recognition*, IEEE Std. 2945-2023, Mar. 2023.

[11] M. Khan *et al.*, "Advances in Facial Emotion Recognition: Deep Learning Approaches and Future Prospects," in *Proc. IEEE Int. Conf. Computer Vision Workshops (ICCVW)*, Paris, France, pp. 450–457, 2024.

[12] R. Patel *et al.*, "Facial Recognition Based on Deep Learning," in *Proc. IEEE Int. Conf. Signal Processing and Computing*, Chennai, India, pp. 67–72, 2024.

[13] S. Kumar *et al.*, "Pre-Trained Deep Learning Models for Facial Emotions Recognition," in *Proc. IEEE Int. Conf. Emerging Computing Technolo- gies*, Singapore, pp. 210–216, 2025.

[14] H. Singh *et al.*, "Deep Learning-Based Real-Time Face Detection and Gender Classification using OpenCV and Inception v3," in *Proc. IEEE Int. Conf. Data Engineering and Applications*, Jaipur, India, pp. 95–101, 2025.

[15] P. Bansal *et al.*, "Age and Gender Prediction from Face Images Using Convolutional Neural Network," in *Proc. IEEE Int. Conf. Image Pro- cessing (ICIP)*, Kuala Lumpur, Malaysia, pp. 3122–3126, 2025.

[16] A. Sharma *et al.*, "ChatGPT and Biometrics: An Assessment of Face Recognition, Gender Detection, and Age Estimation Capabilities," in *Proc. IEEE Conf. Artificial Intelligence and Biometrics*, Tokyo, Japan,
pp. 411–418, 2024.

[17] M. Singh *et al.*, "A Survey of AI-Based Facial Emotion Recognition: Features, ML & DL Techniques, Age-Wise Datasets and Future Di- rections," *IEEE Access*, vol. 12, pp. 12134–12150, May 2024. doi: 10.1109/ACCESS.2024.3366001

[18] X. Li *et al.*, "Efficient Facial Feature Extraction with Hybrid CNNs," in *Proc. IEEE Int. Conf. Computer Vision Systems*, Beijing, China, pp. 199–205, 2024.

[19] Y. Zhang *et al.*, "Transformer-Based Facial Analysis for Real-Time Applications," in *Proc. IEEE/CVPR Workshop on Face Analysis*, San Francisco, CA, USA, pp. 122–129, 2025.